DRAFT

MODELING EMISSIONS DATA SYSTEM

File Structure

Paul D. Allen

April 2, 2001

Control Strategy Modeling Section
Modeling and Meteorology Branch
Technical Support Division
California Air Resources Board

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

Mathematical descriptions of the transport, diffusion, and chemical transformation of air pollutant emissions offer a rational approach to estimating the impact of emission controls on subsequent air quality. Airshed modeling has proven especially valuable in understanding the relationship between ambient levels of ozone and nitrogen dioxide, and their precursor emissions.

The required resolution in the emission inventories has increased correspondingly with the capabilities of airshed models. Inventories are improving both spatially and temporally. The Modeling Emissions Data System (MEDS) contains information necessary for producing 3D photochemical model input emissions files. There is a close association between MEDS and the ARB's emission inventory system used for storage of California emissions data. However, all gridded emissions are stored outside of the statewide inventory system in MEDS formats. MEDS records or files may come directly from ARB's emission inventory system, or from a variety of other sources, such as the DTIM motor vehicle emissions gridding program. Emissions representing specific events (e.g. wildfires) on specific dates are often coded by hand.

Emission data records used in MEDS are stored as formatted sequential files with fixed length ASCII records, to allow easy transport between different computer systems. Emissions in MEDS consist of:

CO	Carbon monoxide
NOx	Nitrogen oxides
SOx	Sulfur oxides
TOG	Total organic gases
PM	Particulate matter

While most of the information contained in MEDS is related to the processing of TOG into specific organic gas species, information on speciating NOx and SOx emissions are also part of MEDS.

EMISSION TRANSACTIONS

Airshed simulation models require hourly gridded emissions data. MEDS is designed to store information needed for pre-processing emissions inventories for inputs to a variety of simulation models. Emission records for all sources are coded in one of two common formats; either MEDS version I or II. Version II formats were

designed to allow hourly specification of variable stack parameters. The emission record formats are shown in Tables 1 and 3.

Facilities with exhaust stacks are handled distinctly. In MEDS version I, the stack data is keyed to the facility, stack, and county codes. The information on the stack record can be used to calculate plume rise. In MEDS version II, the stack keys also include date and time fields. Corresponding stack data formats for point sources are shown in Table 2 and 4.

Emission data are always resolved to one hour time increments. The data usually reflect a typical week day or weekend day. Occasionally, emissions are inventoried for a specific date to enhance air quality model performance evaluations. Emission data records can be coded with a begin hour and an end hour, referenced to a 24 hour clock beginning at 0000 and ending at 2300 PDT. Both begin and end hours refer to the **beginning of the hour**. For example, if the begin hour is 06 and the end hour is 06 the emissions data record refers to only hour 0600-0700. Emission rates on these records are recorded in units of kg/hr. For example, if a source emits 10 kg/hr NOx all 24 hours of the day, the begin hour is 00, end hour is 23, and 10 is coded in the NOx field. Only one emissions data record is needed for this source. If the source emits 10 kg/hr NOx from 0000 to 1200 PDT and 15 kg/hr NOx from 1200 to 0000 PDT then the first record would have a begin hour of 00, an end hour of 11, and 10 in the NOx field. A second record is needed. It would have a 12 for begin hour, a 23 for end hour, and a 15 in the NOx field. (A stack parameter data record may also be needed.)

An alternate method for specifying hourly changes in emission rates for a day is by use of a diurnal profile. The **daily total** emissions in kilograms are recorded on the emissions data record. The begin hour is coded as "-1" and a diurnal profile code is assigned to the end hour. The diurnal profiles currently used are in Table 8. Others may be added as needed.

Table 1

Emissions Data Record Formats

VERSION I

Field	Begin Column	Type. & Length	Description
1	1	A8	Scenario
2	9	I4	SIC
3	13	I8	SCC or CES
4	21	I3	I
5	24	I3	J
6	27	I2	Year
7	29	I2	Month
8	31	I2	Day
9	33	I2	Begin hour or -1
10	35	I2	End hour or diurnal profile code
11	37	I9	Facility ID
12	46	I5	Stack ID
13	51	I2	County
14	53	A3	Air Basin
15	56	F5.0	Elevation (meters above ground)
16	61	F10.1	CO (Kg/hr)
17	71	F10.2	NOx (Kg/hr)
18	81	F10.2	SOx (Kg/hr)
19	91	F10.2	TOG (Kg/hr)
20	101	F10.2	PM (Kg/hr)

Note: If field nine = -1, then field 10 is the diurnal profile code; the emissions are coded as kilograms/day.

File names contain the keyword 'MEDTRANS' or 'MEDS'.

EMISSION TRANSACTION FIELD DESCRIPTIONS - VERSION I

1. Scenario The scenario is a descriptor which will normally be used to differentiate forecasts applicable to the same calendar year. A projected 2010 Sacramento inventory, for example, could be given a scenario name "BASE2010". Controlled 2010 inventories could be named "CTL12010", etc.

2. SIC Standard Industrial Classification codes will be used according to the 1987 edition of "Standard Industrial Classification Manual," Office of Management and Budget, Executive Office of the President of the United States.

3. SCC Source Classification Codes are defined in EPA's AP-42, "Compilation of Air Pollutant Emission Factors." The definitions in ARB's emission inventory system **may not** be the same as those in AP-42, the ARB's MED system's current usage should always be referenced. The SCC code will be used for splitting lumped species on the emissions data record into individual components. The following table shows the splitting available with MEDS:

Emissions	Components
TOG	Total organic gases
NOx	NO, NO_2, nitrous acid
SOx	SO_2, SO_3
PM	Size fractions and chemical species

Many area sources are not defined by SIC/SCC codes. The CES code as used in the ARB's emission inventory system, will be placed in the SCC field and organic and pm species splitting keyed to it. CES codes are 5 bytes and can readily be extracted from SCC records by taking out all SCCs less that 100,000 (integer). This will not separate all area sources from point sources. Some area sources have SCC codes assigned to them, and some point sources have CES codes assigned to them.

4. I Cell number in the X (normally east-west direction.

5. J Cell number in the Y (normally north-south) direction.

6. Year Calendar year the emissions represent.

7. Month Month the emissions represent:

1-12	Data represent the specific month, 1 is January
98	Data represent any month.

81-84	Data represent first, second, third or fourth quarter. January through March is the first quarter.

8. DAY Date the emissions represent:

1-31	Data represent a specific date.
41	Data " a Monday.
42	Data " a Tuesday.
43	Data " a Wednesday.
44	Data " a Thursday.
45	Data " a Friday.
46	Data " a Saturday.
47	Data " a Sunday.
91	Data represent a weekday.
92	Data " a weekend day.
98	Data represent any day.

9. Begin Hour The begin hour is the first hour the data represents. Hour 01 means the emissions from 0100 to 0200.

10. End Hour End hour is the **beginning hour** of the last hour the data represents. If a certain emission record is for the single hour of 10:00 a.m. to 11:00 a.m., the begin hour is 10 and the end hour is 10. Valid hours are 00 through 23 PDT. If a diurnal profile is to be assigned to a source, put a "-1" for begin hour and the 2 digit profile code in the end hour field (see Table 8).

11. Facility ID For most instances, this code will be the same as in the ARB emission inventory. For sources used in forecasts only, the code can be assigned any number that does not conflict with codes already in use.

12. Stack ID For facilities with more than one stack, this ID code allows identification of each separate stack.

13. County A numeric code is assigned alphabetically to each county. See Table 5 for county code definitions.

14. Air Basin An alpha abbreviation is assigned to each air basin, e.g. SV for Sacramento Valley. See Table 6 for air basin code definitions.

15. Elevation This information (in meters above surface) is used when emissions from such sources as aircraft are included in the inventory.

16. CO (Kg/hr) Hourly emissions of carbon monoxide, unless begin hour = -1, then Kg/day.

17. NOx (Kg/hr) Hourly emissions of oxides of nitrogen (as $N0_2$), unless begin hour = -1, then Kg/day.

18. SOx (Kg/hr) Hourly emissions of sulfur oxides (as $S0_3$), unless begin hour = -1, then Kg/day.

19. TOG (Kg/hr) Hourly emissions of total organic gases, unless begin hour = -1, then Kg/day. Many variations exist in molecular weights which are assigned in calculating TOG mass emissions. Corrections are made in the profiles assigned to each emission category. Do not put NMHC, RHC, or ROG in this field.

20. PM (Kg/hr) Hourly emissions of particulate matter, unless begin hour = -1, then Kg/day.

Table 2

Stack Parameter Data Record Formats

VERSION I

Field	Begin Column	Type & Length	Description
1	1	I9	Facility ID
2	10	I5	Stack ID
3	15	I2	County
4	17	F9.3	UTM-east (Km)
5	26	F9.3	UTM-north (Km)
6	35	F8.0	Stack Height (meters)
7	43	F8.1	Diameter (meters)
8	51	F8.1	Gas velocity (meters/second)
9	59	F8.0	Exit temperature (Kelvin)
10	67	F8.1	Flow Rate (m3/second)

Note: Fields 1 through 3 combined (16 bytes), form the key to the emissions data records.

File names may have the keyword 'MEDTRANS' or 'MEDS', but always contain the word 'STACK'.

UTM zone to be added as an 11th field in next revision of MEDS.

STACK PARAMETER DATA FIELD DESCRIPTIONS - Version I

1. Facility ID A code assigned to each facility in EDS. Facility names can be extracted from EDS using RAMIS. Facility IDs can be repeated for different counties.

2. Stack ID The stack number within a facility, also contained in EDS.

3. County These first three fields are the "keys" to identify stack data with its emission data. The county code is a 2 byte integer field and is the same as shown in Table 5.

4. UTM East Universal Transverse Mercator coordinates of the stack, (east-west kilometers). UTM zone should be consistent with the zone used for modeling domain definition.

5. UTM North Universal Transverse Mercator coordinates of the stack, (north-south kilometers).

6. Stack Height The height of the stack measured from the ground surface (meters).

7. Diameter The inside diameter at the top of the stack (meters).

8. Gas Velocity The exit gas velocity, (meters/second).

9. Exit Temperature The temperature of the exhaust upon exit, (degrees Kelvin).

10. Flow Rate Volume flow of exhaust. This may be omitted if gas velocity and diameter are known. However, always include diameter and either the exhaust velocity or the flow rate, ($meters^3$/second). This flow rate is the actual flow rate at the temperature of the exhaust gases. Do not convert to a standard temperature or pressure.

Table 3

Emissions Data Record Formats

Version II

Field	Begin Column	Type. & Length	Description
1	1	A8	Scenario
2	9	I4	SIC
3	13	I8	SCC or CES
4	21	I3	I
5	24	I3	J
6	27	I9	Facility ID
7	36	I5	Stack ID
8	41	I2	County
9	43	I5	Julian day
10	48	I2	Begin hour or -1
11	50	I2	End hour or diurnal profilecode
12	52	A3	Air Basin
13	55	A3	Sub County
14	58	F5.0	Elevation (meters above ground)
14	63	F10.1	CO (Kg/hr)
15	73	F10.2	NOx (Kg/hr)
16	83	F10.2	SOx (Kg/hr)
17	93	F10.2	TOG (Kg/hr)
18	103	F10.2	PM (Kg/hr)

Note: If field 10 = -1, then field 11 is the diurnal profile code. The emissions are coded as kilograms/day. Always include decimals in “F” fields.

File names contain the keyword 'MEDTRANS' or 'MEDS'.

EMISSION TRANSACTION FIELD DESCRIPTIONS - VERSION II

Field descriptions are the same as version I, except for the following fields.

9. Julian Day YYDDD (87238) Julian day the emissions represent. The Julian day is normally used only for data which represents a specific day. This field is not normally used in processing emissions since an emissions file cannot contain emissions for more than one day. The normal use is to use the date in the file name. However, the Julian day and hour fields are part of the stack key when hourly stack parameter information is available. A Julian day of 400 is used for weekday. A Julian of 500 is used for weekend days. Use 00 for year 2000, 10 for 2010, etc.

13. Sub County This field is used for preparing reports from emission records. An example is Kern. Emissions control regulations require different controls for western Kern than central Kern. This field is left blank if the county is not divided.

County	Sub County Code	Description
Kern	WST	Western Kern
	CNT	Central Kern
Santa Barbara	NOC	North County
	SOC	South County
	L/S	Lompoc/Santa Ynez
	OCS	Outer Continental Shelf

EMISSION TRANSACTION FIELD DESCRIPTIONS - VERSION III

Field descriptions are the same as version II, except for the following fields.

2. SIC The 14 byte CEIDARS SIC field replaces the traditional 4 byte SIC code.

3. SCC The 14 byte CEIDARS SCC field replaces the traditional 8 byte CES or SCC fields.

Table 4

Stack Parameter Data Record Formats
VERSION II

Field	Begin Column	Type & Length	Description
1	1	I9	Facility ID
2	10	I5	Stack ID
3	15	I2	County
4	17	I5	Date (Julian)
5	22	I2	Begin hour
6	24	I2	End hour
7	26	F9.3	UTM-east (Km)
8	35	F9.3	UTM-north (Km)
9	44	I2	UTM zone
10	46	F8.0	Stack Height (meters)
11	54	F8.1	Diameter (meters)
12	62	F8.1	Gas velocity (meters/second)
13	70	F8.0	Exit temperature (Kelvin)
14	78	F8.1	Rate (m /second)

Note: Fields 1 through 8 combined (31 bytes), form the key to the emissions data records, if stack parameters vary with time or location (large ships) these keys will allow matching to the emission data records.

File names may have the keyword 'MEDTRANS' or 'MEDS', but always contain the word 'STACK'.

STACK PARAMETER DATA FIELD DESCRIPTIONS VERSION II

Descriptions are the same as version I except for the following:

4. Julian Day

5. Begin Hour Included in stack record to allow hourly variations in stack parameters.

6. End Hour

9. UTM Zone Always use zone consistent with modeling region definition.

TABLE 5

COUNTY CODES AND NAMES

Code	County	Code	County
01	Alameda	30	Orange
02	Alpine	31	Placer
03	Amador	32	Plumas
04	Butte	33	Riverside
05	Calaveras	34	Sacramento
06	Colusa	35	San Benito
07	Contra Costa	36	San Bernardino
08	Del Norte	37	San Diego
09	El Dorado	38	San Francisco
10	Fresno	39	San Joaquin
11	Glenn	40	San Luis Obispo
12	Humboldt	41	San Mateo
13	Imperial	42	Santa Barbara
14	Inyo	43	Santa Clara
15	Kern	44	Santa Cruz
16	Kings	45	Shasta
17	Lake	46	Sierra
18	Lassen	47	Siskiyou
19	Los Angeles	48	Solano
20	Madera	49	Sonoma
21	Marin	50	Stanislaus
22	Mariposa	51	Sutter
23	Mendocino	52	Tehama
24	Merced	53	Trinity
25	Modoc	54	Tulare
26	Mono	55	Tuolumne
27	Monterey	56	Ventura
28	Napa	57	Yolo
29	Nevada	58	Yuba

TABLE 6

CALIFORNIA AIR BASINS

Code	Air Basin
NC	North Coast
SF	San Francisco Bay Area
NCC	North Central Coast
SCC	South Central Coast
SC	South Coast
SD	San Diego
NEP	Northeast Plateau
SV	Sacramento Valley
SJV	San Joaquin Valley
GBV	Great Basin Valleys
MD	Mojave Desert
SS	Salton Sea
MC	Mountain Counties
LC	Lake County
LT	Lake Tahoe
SED	Southeast Desert (*historic*)

DIURNAL DISTRIBUTIONS

Diurnal (or temporal) distributions can be used to define hourly emission variations during a day when appropriate. Certain emission categories may display consistent patterns that may be measured or estimated to give a more realistic diurnal profile of pollutant releases than assuming emission rates are constant during a day. The time of release can be important for most types of modeling studies. Table 8 shows weight fractions by hour for each diurnal distribution code used in MEDS. This code is found in the 'END HOUR' field on the emission record when the 'BEGIN HOUR' is set to -1. Any temporal profile can be used by simply specifying hourly emissions for a given source, that is, set 'BEGIN HOUR' equal to 'END HOUR' and code in that hour's emissions. This requires 24 separate emission records to make up one entire day, but is very useful for sources like power generating facilities which have large emissions and also maintain records of hourly fuel usage. The profile codes have the following usage in MEDS, but can be used for any source with similar hourly distributions (partial listing):

Code	Description
31	Gasoline service stations
33	Residential fuel combustion
34	Orchard heaters
35	Commercial aircraft
37	Daylight hours (summer)
38	Residential cooking
50	Motor vehicle usage
51	Petroleum dry cleaning
52	Pesticide application (Ventura county)
53	Agricultural aircraft activity (Ventura county)
54	Daytime biogenic emissions
55	Nighttime biogenic emissions

Emissions records are always referenced to begin hour in Pacific Daylight Time (PDT). Table 7 shows the formats for the temporal distribution files. Conversion to PST is best handled for an entire episode after preprocessing for a specific model. It is conventional to reference times in all modeling files to begin hour in PDT (hour 00 represents 0000 to 0100)).

Table 7

DIURNAL DISTRIBUTION FILE FORMATS

Field	Type & Length	Description
1	I2	Diurnal distribution code
2 - 25	F6.4	24 hourly fractions

There is one MEDS files for diurnal distributions.

pdt.d : Record length = 146

ORGANIC GAS SPECIES PROFILES

Several related files are needed to resolve total organic gas (TOG) emissions into individual organic species and subsequently to lumped species needed for reports or airshed models. Table 8 shows the file format for assigning emission categories to organic gas profiles. The organic gas profiles are keyed to the "SCC" or “CES”fields.

Table 8

SOURCE CATEGORY DESCRIPTION FILE

Field	Type & Length	Description
1	I8	Source category (SCC or CES)
2	I4	Organic profile code
3	A70	Source category name

Filename is sccnames.* : Record Length = 100

The organic gas profiles have been compiled from many sources, and research to update profiles is continuous. Appendix A contains documentation for each profile. The formats for the emission profiles of organic gases are shown in Table 9. These profiles contain weight fractions for each chemical. There are two files containing organic gas weight fractions MEDS. The first has profiles as reported from laboratory analyses, the sum of species should total to 100%. It is often necessary to correct these profiles to account for an assumed molecular weight used in the emission estimate. For instance, for some categories of combustion, mass emissions are calculated by using a molecular weight of 16 (methane) for each carbon measured. Emissions are then reported "as methane". The weight fractions for each species must be adjusted to account for the assumed molecular weight. In some cases, emissions are estimated as total hydrocarbons (THC) and do not include oxygenated compounds. The profiles are adjusted to "add" these compounds where appropriate. In most cases the corrected weight fractions for combustion sources do **not** total exactly to 100 per cent.

TABLE 9

ORGANIC GAS PROFILE DATA

Field	Type & Length		Description
	1.	2.	
1	I5	I5	Organic profile code
2	I5	I6	SAROAD chemical code
3	F8.4	F7.2	See note below

1. **Corrected** profiles are in
   outpro4.d
   Record length = 18
   Field 3 is **weight fraction**

2. **Uncorrected** profiles are in
   profiles.d
   Record length = 80
   Field 3 is **weight percent**

The information contained in Tables 8 and 9 are used to speciate TOG into individual organic gas compounds. An organic gas species profile is often assigned to many different emission categories. The profiles are developed from source tests, literature review, or in some cases - engineering estimates. Associated with combustion profiles are the NO and NO_2 splits, and the SO_2 and SO_3 splits from the inventoried NOx and SOx emissions. At the present, the NOx and SOx splits are carried on the profile names file (pnames.d). Table 11 shows the formats for this file. It is important to note that NOx emissions are estimated using the molecular weight of NO_2, while SOx emissions are estimated as SO_2. The NOx and SOx fractions in MEDS can be varied by profile, but at present are constant for all sources. NO is 88% of the NOx, NO2 is 10%, and nitrous acid is 2%. SO2 is 97% of the SOx while SO3 is the remaining 3%.

TABLE 10

ORGANIC GAS PROFILE NAMES

Field	Length	Description
1	I5	Organic profile code
2	A72	Profile name
3	F6.2	Assumed molecular weight
4	F6.2	NOx to NO fraction
5	F6.2	NOx to $N0_2$ fraction
6	F6.2	SOx to $S0_2$ fraction
7	F6.2	SOx to $S0_3$ fraction
8	I2	0 if emissions are TOG 1 if emissions are THC

Filename is pnames.d : Record length = 120

Information on the organic gas compounds is kept in the chemical file which is defined in Table 11. The key field is the chemical code. The Chemical Abstract Service (CAS) code is being added to the chemical reference file but is not complete. To avoid errors, always refer to the 5 byte chemical code and definitions as contained in the MEDS chemical file when using the MEDS organic gas species profiles.

Note that when using this file, many entries are for mixtures. ARB staff is preparing a 'mixture' definition file which can be used to completely eliminate mixtures from the subsequent organic gas inventory.

Field 7 contains a 1 for all species designated as part of ROG by ARB's Emission Inventory Branch. Species excluded from ROG are designated with a 2

Table 11

ORGANIC GAS CHEMICAL DATA

Field	Length	Description
1	I5, 1x	Chemical code
2	A12, 1x	CAS Registry number
3	A65	Chemical name
4	F5	Carbon number
5	I5	Organic gas group
6	F10.2	Molecular weight
7	I1	ARB ROG class (1=ROG)
8	F12.5	O rate constant
9	F12.2	OH rate constant
10	F12.8	O3 rate constant
11	*A16*	*Formula (to be added)*
12	i10	CAS without the dashes
13	A1,2x	Y is CAS is for ARB use only
14	A32	Chemical name used before 1998

Rate constant units are 1/(ppm*min)
Filename is chem3.d

Table 12

Organic Gas Group Definitions

(Field 5)

1	Low reactives	13	Halogens	25	Propylene
2	Ethylene	14	Terpenes	26	1,3-butadiene
3	Benzene	15	Glycols	27	Toluene
4	C6+ Alkanes	16	Styrenes	28	Acetaldehyde
5	C4+ Alkenes	17	Alkynes	29	MTBE
6	C8+ Aromatics	18	Amines	30	Ethanol
7	C3+ Aldehydes	19	Formaldehyde	31	Acetylene
8	Alcohols	20	Methane	32	Isoprene
9	Ketones	21	Ethane	98	Unclassified
10	Esters	22	Propane	99	Unidentified
11	Ethers	23	Butanes		
12	Acids	24	Pentanes		